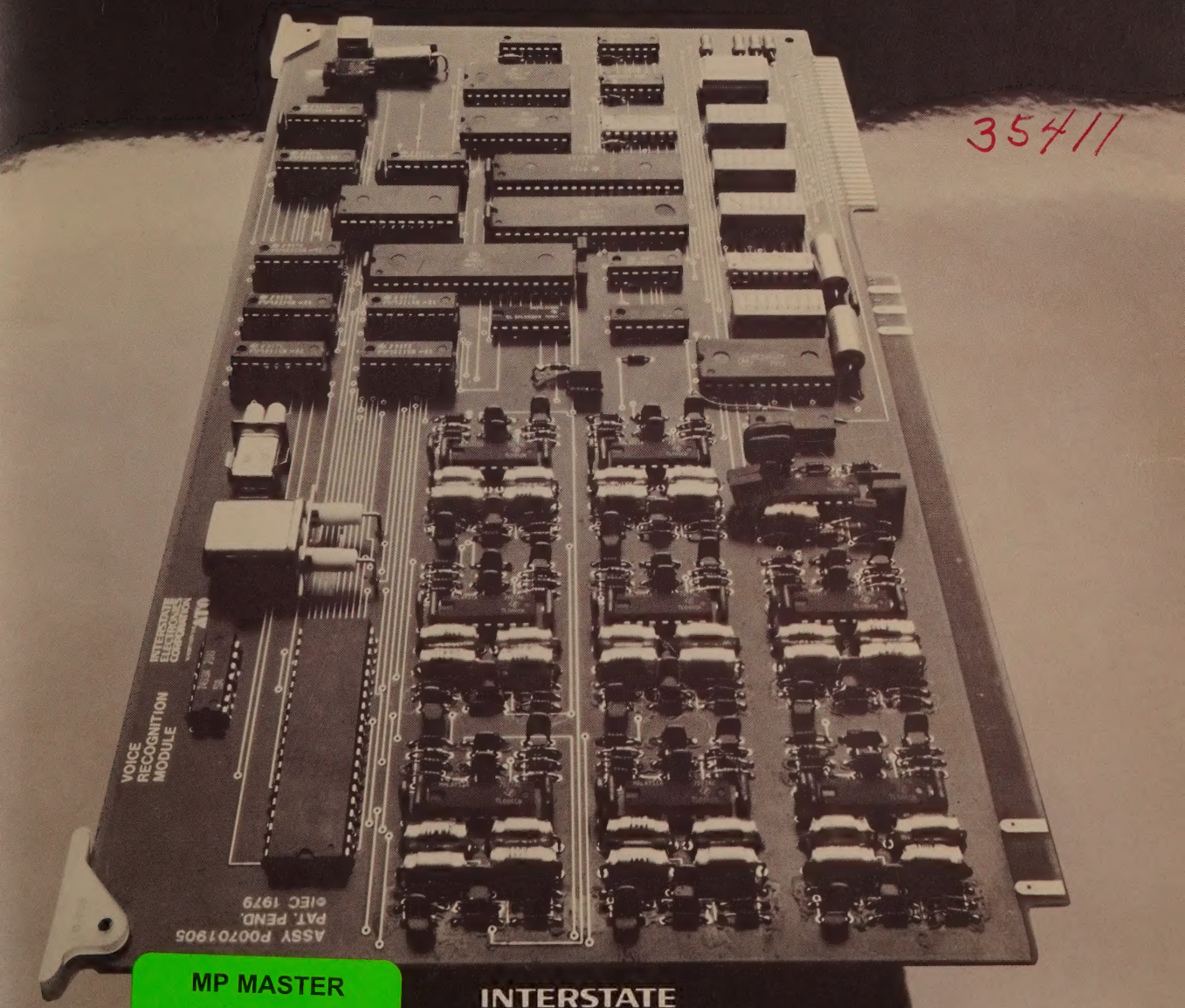
VOICE RECOGNITION MODULE (VRM) REFERENCE MANUAL (Includes Voterm)



35411

MP MASTER
MA28300

INTERSTATE ELECTRONICS CORPORATION
SUBSIDIARY OF ATO

VOICE RECOGNITION MODULE (VRM) REFERENCE MANUAL (Includes Voterm)

Revision 3 – April 1981

This manual pertains to those VRM's effected by IEC drawing POO 701905 dated October 28, 1980. In addition, it will apply to any VRM's that have been modified by IEC to incorporate this ECO.

For service and manufacturing information call (714) 635-7210, extension 6493.

INTERSTATE ELECTRONICS CORPORATION
SUBSIDIARY OF ATO

CONTENTS

Paragraph		Page
	Chapter 1. Introduction	
1.0	Introduction	1
	Chapter 2. Description	
2.0	Description	3
2.1	VRM Simplified Functional Diagram	3
2.1.1	Port A – Parallel Host Communications	3
2.1.2	Port 1 – Serial Host Communications	4
2.1.3	Port 2 – Auxiliary Serial Communications	4
2.1.4	Port B – User Control	5
2.1.5	Audio Input	5
2.2	Description of Recognition System Specification	5
2.2.1	Recognition Performance	5
2.2.2	Reject Threshold	7
2.2.3	System Response	7
2.3	Programming	8
2.3.1	Basic Host-to-VRM Commands	8
2.3.2	Extended Host-to-VRM Commands	9
2.3.3	Command Summary	10
2.3.4	VRM System Compatibility Command	10
2.3.5	Programming Details, VRM/Host Commands	13
2.4	Electrical and Interface Requirements	26
2.4.1	Power Requirements	26
2.4.2	Audio Input	27
2.4.3	Digital Input/Output	27
2.4.4	Data/Command Synchronization	27
2.4.5	Error Conditions	30
2.4.6	System Configuration	30
2.5	Mechanical	31

Paragraph		Page
2.5.1	Board Size	31
2.5.2	Connectors	31
2.5.3	Card Cage	31

Chapter 3. Voterm Hardware Description

3.1	Installing a VRM	37
3.2	Reset Switch	40
3.3	Connector Pin Assignments	40
3.4	Air Filter Maintenance	40
3.4.1	Visual Inspection	40
3.4.2	Removing Filter	40
3.4.3	Filter Assembly	40
3.4.4	Cleaning the Filter	40

FIGURES

Figure		Page
	Chapter 2. Description	
2-1	VRM Simplified Functional Diagram	4
2-2	VRM Interface Reference	6
2-3	Parallel Input Handshaking (Host to VRM)	28
2-4	Parallel Output Handshaking (VRM to Host)	29
2-5	VRM Switch Locations	30
2-6	VRM Board Outline	35
	Chapter 3. Voterm Hardware Description	
3-1	Front Panel of Voterm II (With Shure SM-10 Microphone)	38
3-2	Top View of Voterm II, Showing Meter Drive PCB	38
3-3	Top View of Voterm II, Showing 110-220 Vac Power Switch	39
3-4	Rear Panel of Voterm II	39

TABLES

Table		Page
	Chapter 2. Description	
2-1	Command Summary	11
2-2	Function vs. Switch/Logic Setting	32
2-3	VRM Power-Pin Assignments	34
2-4	Signal Connections for Host Mode	34
2-5	VRM Specification Summary	36
	Chapter 3. Voterm Hardware Description	
3-1	Parallel-Port Connections for Input/Output	41
3-2	Parallel-Port Connections for Configuration/Control	41
3-3	Serial-Port Connections to Host Processor	42
3-4	Serial-Port Connections to Peripheral Device	42
3-5	Available Voterm DC Power	42

Chapter 1

INTRODUCTION

1.0 INTRODUCTION

The Voice Recognition Module (VRM), a product of Interstate Electronics Corporation, is a single printed-circuit board automatic speech recognizer consisting of a CPU, 4K bytes of ROM, up to 4K bytes of RAM for reference pattern storage, analog circuitry for speech spectrum analysis, electronics for interfacing to a controlling host processor, and provisions for optional serial interfaces. Communications to the controlling host processor can be either parallel or asynchronous serial (first optional serial port). The VRM recognizes isolated utterances. Each recognized incoming utterance is compared to a reference pattern obtained by initially training each user.

The original VRM was introduced by Interstate Electronics Corporation in November 1979. Since its introduction, an extended set of host-to-VRM commands have been incorporated which enhance the utility and versatility of the original version. In addition, high speed logic (microprocessor and I/O) have been added to improve the voice processing and data transfer times. The current VRM is 100 percent compatible with all host computer software developed for the earlier VRM except the change control character command. It retains all nine of the original host-to-VRM commands while adding an additional five commands. The current VRM, unlike its predecessor, does not support stand-alone operation; it is exclusively controlled via a host processor.

The VRM allows the vocabulary size, recognition decision threshold, training mode, and number of training passes to be selected by the user. Individual vocabulary items or vocabulary subsets can be updated by additional training examples or replaced by new ones. Because any contiguous subset of vocabulary items is selectable in the recognition mode, the VRM has a selectable syntax structure.

Additionally, a second contiguous vocabulary subset can be selected and appended to any primary subset, allowing the use of "common vocabulary" as typically associated with a control vocabulary.

Finally the new feature, the extended syntax recognition command, allows total random vocabulary syntax selection.

The board functions are completely controlled by the host processor. Where long line communication is necessary, asynchronous serial communication provides minimal interconnections and

good noise immunity. Where the VRM is to be housed in the same chassis, or near-by the host (15 feet or less), the standard parallel interface affords higher speed communication.

The vocabulary reference patterns can be transferred either from the host processor to the VRM or from the VRM to the host for permanent storage of different speaker reference patterns, or for use with the VRM for vocabularies exceeding 100 words.

Since the specific users application program resides in the host, a highly interactive man-machine communications system can be implemented using man's speech as a means for control, inquiry, or on-line data acquisition at a relatively low cost per user. All real-time speech signal preprocessing and recognition are performed by the VRM, thereby allowing the host processor maximum time to perform the application task. This feature allows even a microprocessor host to operate on multiple recognition responses of several VRM's while allowing ample time for execution of complex application programs in real-time.

Chapter 2

DESCRIPTION

2.0 DESCRIPTION

This chapter describes the VRM in terms of a simplified functional block diagram, and provides additional clarification of the recognition performance, reject threshold, and the system response time specification (as listed at the end of this document). It further summarizes and describes in detail the host-to-VRM commands, and covers the mechanical and electrical characteristics of the VRM.

2.1 VRM SIMPLIFIED FUNCTIONAL DIAGRAM

A simplified functional block diagram of the VRM is shown by figure 2-1. The decision whether to utilize the parallel interface to the host or the optional serial interface must be made based on several considerations, including; 1) distance between the VRM and the host machine, 2) serial baud rates available at the host, 3) required upload/download response times, and 4) ease of implementation. At distances greater than four feet, timing problems preclude parallel interfacing to the host. If only 110 or 300 baud are available (typical for hard-copy terminals), then generally these slow rates tend to cause a perceivable deterioration in recognition response time, even though the VRM is relatively fast (see section 2.2.3.3) in burst recognition. The host is frequently called upon to perform additional tasks. Timings for such host tasks can make the slower serial transfer rates very attractive by comparison. Because the completed execution of any application task occurs during recognition, the decision whether parallel or serial interfacing is required must be based primarily on the complexity and speed of further host actions that will be taken during recognition.

All data transmitted to or from the host is ASCII data, except when the high-speed parallel transfer flag is set and an upload or download of reference patterns is requested.

2.1.1 Port A – Parallel Host Communications

This parallel port provides eight data input lines for receiving commands and downloading speech reference patterns; eight data output lines for training prompts, recognition output in response to an utterance, reference pattern uploading to the host, and acknowledging all commands. Each character received from the host is accepted in less than 160 microseconds. Outputs from the VRM to the host require less than 120 microseconds. Four handshaking lines for parallel data transfer are provided. Two lines are associated with data input, two with data output. Pulse or level handshaking modes may be selected by the user via on-board switches or TTL levels.

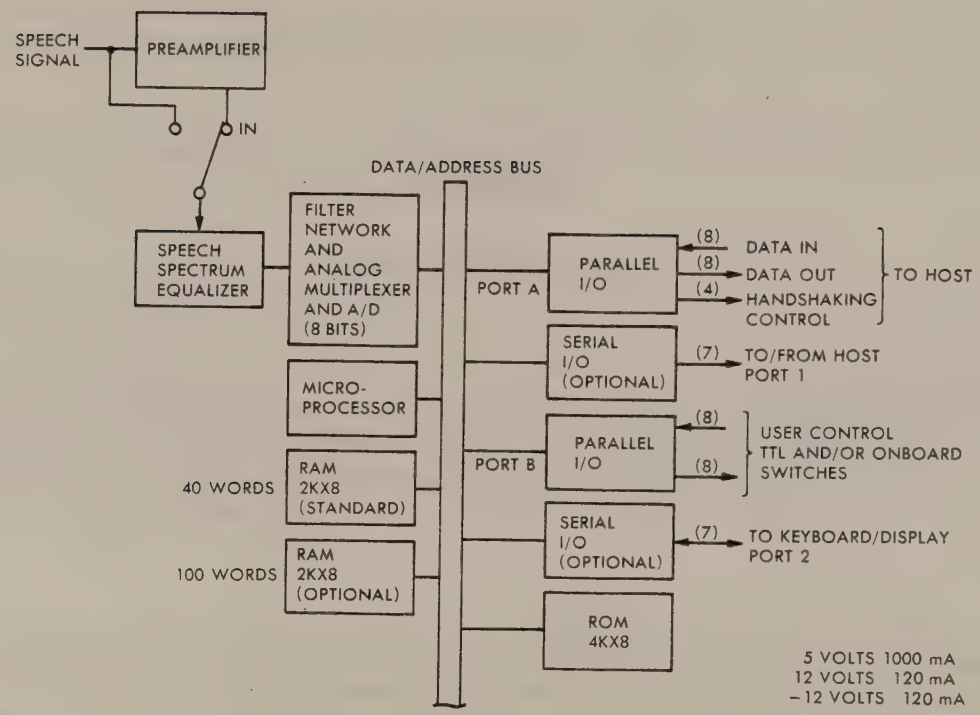


Figure 2-1. VRM Simplified Functional Diagram

2.1.2 Port 1 – Serial Host Communications

This optional serial asynchronous port allows RS232C or current loop communication at user-selected rates from 110 to 9600 baud. The user may select, via on-board switches or TTL levels, the serial word data format, i.e., number of stop bits, parity and 7-or 8-data bits. Serial port host handshaking, if required, may be accomplished by host control of the VRM clear-to-send output, CTS, associated with this port, using RS232C logic levels. The VRM is always ready to receive host commands or data; hence, the request-to-send line, RTS, is always high (+12 volt D.C.). The RS232C or current loop option is selected via an on-board switch.

2.1.3 Port 2 – Auxiliary Serial Communications

In order to provide asynchronous communication to an auxiliary terminal, display, printer, etc., the VRM has provisions for a second optional serial interface. The interface operates at the same baud rate and serial word format as selected by the user for Port 1. The RS232C or current loop option is selected via an on-board switch. Port 1 and Port 2 RS232C or current loop may be selected to be the same or differing, since these are independent switches.

Port 2 allows a host to communicate with the VRM and an auxiliary serial device while using only one serial host communication channel.

Port 2 is a two-way communication path; i.e., ASCII input to this port is passed directly, to the host processor. If keyboard data is entered via this port, remote commands from the users work-site can be sent to the host directly.

2.1.4. Port B – User Control

The user may control via on-board switches SA, SB, and SC various system configurations. Switches SB and SC, or their equivalent TTL logic levels, are read on power-up, or upon a manual or software reset. Switch SC selects parallel or serial mode, echo/no echo, carriage return or carriage return and line feed, parallel host-handshaking mode, and power-up/reset acknowledge. Switch SB selects the serial word format (three positions). Switch SA selects serial baud rate.

2.1.5. Audio Input

The VRM includes an on-board microphone preamplifier and a two-position preamplifier bypass switch (switch E on the VRM board). The preamplifier IN position on this switch allows direct microphone input using boom microphones such as the Shure Brothers SM-10, Telex CS-75, or equivalent microphone. The VRM is relatively independent of the speech-energy level of each isolated utterance for a typical speaker. However, to accommodate variations in speaking level, resistor R114 of the on-board preamplifier may be reduced for speakers with a weak vocal effort or increased for a strong speaker (see VRM Interface Reference drawing for details, figure 2-2).

Alternately, an audio signal from an external microphone amplifier can be used by bypassing the on-board amplifier (switch E set to the OUT position). This feature accommodates use of a remote microphone and preamplifier without losing audio signal integrity. The external preamplifier should, on the average, have an output of 6.5 $\pm$2 volts peak-to-peak maximum for each isolated input utterance. This provides nearly optimal VRM recognition. The VRM useful audio bandwidth is 200 to 7000 Hz. Excellent recognition is attainable with reduced telephone bandwidths.

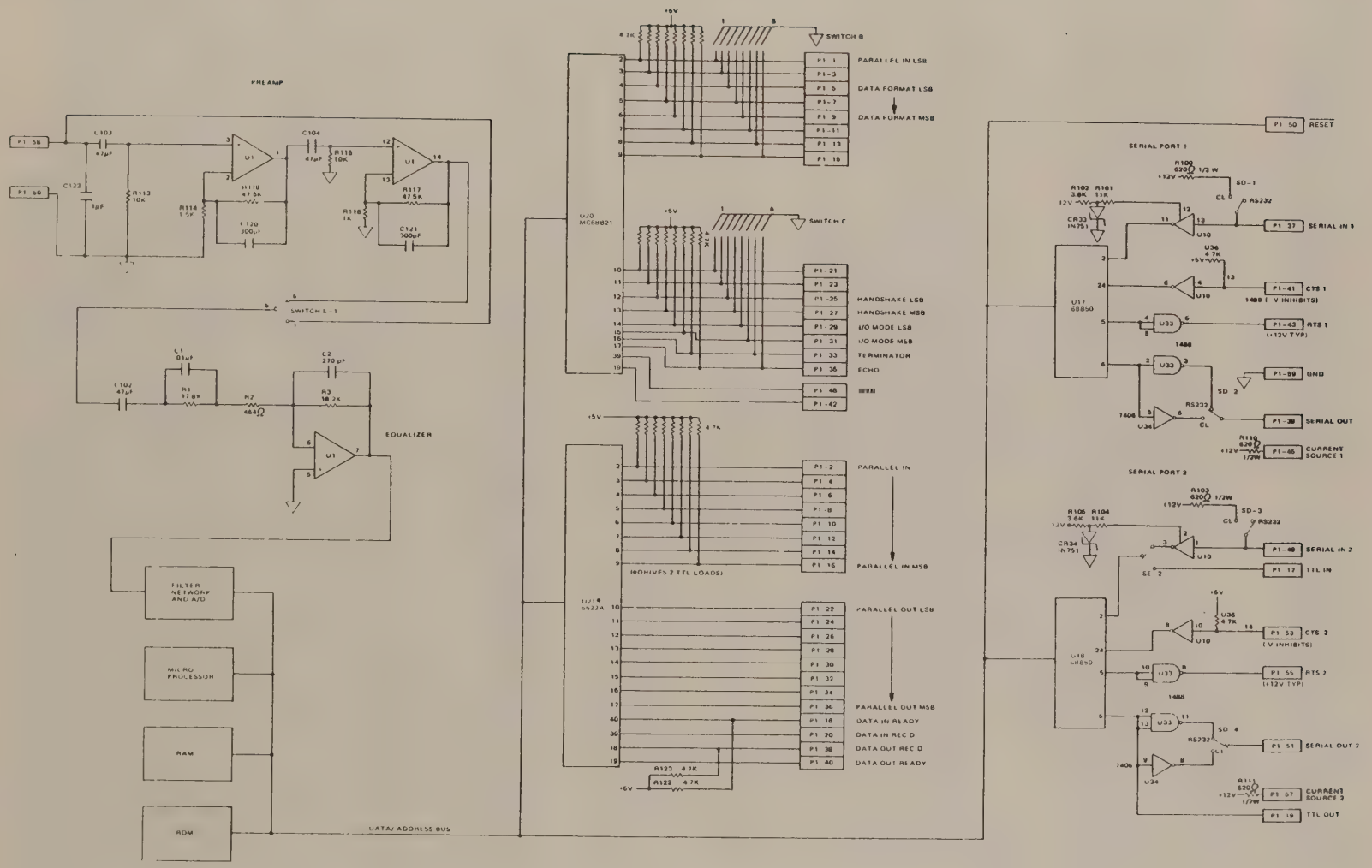
2.2 DESCRIPTION OF RECOGNITION SYSTEM SPECIFICATION

This section decribes and clarifies the recognition performance, reject threshold, and system response times of the VRM.

2.2.1 Recognition Performance

The VRM has been tested by a number of speakers with three vocabularies: (1) a vocabulary consisting of the ten digits and resulting in a measured average accuracy of 99.7 percent; (2) a 42-word vocabulary of ten digits, International Civilian Aeronautic Organization (ICAO), and six other noncommand words with an average accuracy of 99.3 percent; and (3) a 100-word test with an average accuracy of 95.8 percent.

The 100-word test included 50 phonetically balanced words found in ANSI S3.2-1960 (R1971), plus the digits, the ICAO alphabet (e.g., ALPHA, BRAVO, CHARLIE . . . ZULU), and 13 control words (e.g., LINE FEED, RETURN, KILL, ESCAPE, etc.). This test is extremely difficult for a recognition system because it includes several similar-sounding words that differ only by a single phoneme (e.g., GROVE/CLOVE, MIKE/DIKE, FOUR/FORD, ARE/BAR, etc.). This specific vocabulary of 50 phonetically similar words is not a typical-application vocabulary requirement. In any application vocabulary, homonyms (e.g., for/four, one/won, two/to) should be avoided.



7902-966-001A

Figure 2-2. VRM Interface Reference

Recognition accuracy is a function of training examples. Optimum performance is generally achieved after seven to ten training examples of each word or phrase. The specified recognition accuracy of 99 + reflects the percentage of times the classification result will be correct on the first attempt when a valid utterance is spoken. This specified accuracy value can be attained with properly trained speakers and typical application vocabularies of up to 100 words in an 80 dBm acoustic environment.

2.2.2 Reject Threshold

The reject threshold is user-selectable between 0 and 128 in unit increments. The threshold is used to compare the degree of precision in the match between the input utterance pattern and the reference pattern. A higher value results in a better rejection of invalid words, invalid speakers, and background noise at the expense of a slightly greater frequency of rejection of valid words.

2.2.3 System Response

Four significant VRM duration and response time parameters are described in the following paragraphs.

2.2.3.1 Minimum Between-Word Pause (ETHL parameter) – The VRM is a discrete-word recognizer. Many discrete words have a silence pause. Examples of such words are "abort", "token", "eight", "stop", "help", and "delete". Each of these example words has a silence interval caused by one or more stop sounds (b, d, g, k, p, t). The VRM is also capable of recognition of short phrases and polysyllabic words. Discrete word/phrase boundaries are detected based upon the spectral energy of the set of VRM filters first exceeding, and then returning, to an ambient silence background level. For this reason, an operator must pause between utterances. A nominal 160-millisecond minimum pause requirement ensures that naturally occuring gaps in discrete utterances will not inadvertently be confused with between-utterance pauses. The 160-millisecond pause will occur naturally if the operator simply "thinks" of pausing following each utterance. This parameter, however, may be set to a value differing from the default value by a host-to-VRM command.

2.2.3.2 Min/Max Word Length – The VRM accepts a word or phrase whose duration is from 80 to at least 1250 milliseconds. The minimum word length assists the VRM in discarding abrupt impulse noise, such as the dropping of a tool or brushing against the microphone. This parameter (MINSM) may be set to a value, differing from the default value, 80 milliseconds, by a host-to-VRM command.

The VRM saves a maximum of 250 significant samples per utterance. A new sample is taken every 5 milliseconds, but if it is not significantly different from the previous sample, it is discarded. Assuming uniform samples (all samples significant), the longest allowable utterance is 5X250=1250 milliseconds. Assuming all samples are significant depends upon the value of T_2. Increasing T_2 beyond the default value of 16 ensures that utterances longer than 1.25 seconds can be accepted.

A 1.25-second allowable utterance duration permits entry of short phrases. It also enables the VRM to discard continuous speech as spurious input; thus, if the operator converses with another person while wearing a live VRM input microphone, such conversation will not result in a series of invalid recognition responses.

2.2.3.3 Response Time – The VRM response time consists of two parts: (1) the between-word pause duration, and (2) detected utterance classification processing time. Thus,

RESPONSE TIME = PAUSE + PROCESSING TIME

The default pause duration (ETHL) is 160 milliseconds. The processing time to determine the best match of the input pattern with each reference pattern in the active vocabulary (recognition plus common vocabulary) may be expressed by the formula:

PROCESSING TIME = (25+N) milliseconds

where:

N = active number of reference patterns to be compared with the incoming pattern

hence,

RESPONSE TIME = ETHL + (25+N) milliseconds

If the default value for ETHL = 160 milliseconds is substituted, then,

response time = 185 + N milliseconds

Assuming a vocabulary such as the digits, the VRM provides burst recognition of about 90 words/min. (90 word/min = 1.5 words/sec or 667 msec/word).

2.3 PROGRAMMING

2.3.1 Basic Host-to-VRM Commands

VRM control is effected by host commands via the I/O required (standard parallel, or optional serial ports). The system accepts 14 commands; nine of the these commands are identical with those of the original VRM.

With the basic nine commands, all information flow between the host processor and VRM – regardless of direction – is in the form of ASCII characters whether the transfer is parallel or serial, except when a user selects high-speed parallel transfer and an upload or download command is issued. For reference pattern transfer, bits 7, 6, and 5 are always 011. Hence, reference patterns yield ASCII characters of 0,1,2,3,···8,9,:,;,<,=,>,?,– never a control character. The VRM inserts this coding to ensure that the host does not interpret data as control characters, and the host can store and output reference patterns directly in ASCII code form.

The nine basic commands are: TRAIN, UPDATE, RESET, SET REJECT THRESHOLD, PORT 2 TRANSFER, DOWNLOAD, UPLOAD, SELECT COMMON VOCABULARY and RECOGNIZE.

Host commands for the VRM or outputs from the VRM in response to completion of the command function, or in response to other events, adhere to the following general format:

FRAMING CHARACTERS, n, CHARACTER STRING, TERMINATOR

where:

n = command identifier, "1" to "9" or "A" through "E", indicating host command or a VRM response to the command.

CHARACTER STRING = control parameters or text

FRAMING CHARACTERS/CHARACTER = ASCII control characters DC1, DC2, DC3, or DC4, unless specified differently by the user following a power-up.

TERMINATOR = message/command terminator in the form of a carriage return (CR) or carriage return and linefeed (CR,LF), unless otherwise specified following a power-up.

The nine basic commands use only the digits 1 through 9 for the command identifier. They are detailed in section 2.3.5. The five extended commands are summarized below, and detailed in section 2.3.5 as well.

2.3.2 Extended Host-to-VRM Commands

Five of the six extended commands of the current VRM use only one letter (A through E) for the command identifier.

These commands are:

Command Identifier Letter	Summary Description
A	Set/Clear Mode flag. An ASCII "1" sets a specific mode while an ASCII "0" clears or resets the mode. The six mode flags in their sequential order are: (1) Extend data available during recognition (2) Acknowledge-all-host commands (3) High-speed parallel upload and download reference patterns (4) Extend vocabulary size beyond 100 words (not currently available).

(5) Transmit all noncommand characters from host to port 2.
(6) If greater than 250 significant samples, this flag causes "LL" to be output in place of the word number in the recognition response for utterances that are too long.

B — Read-reject threshold.

C — Write-new-word boundary parameters (T_1, T_2, MINSM, ETHL).

D — Read-word boundary parameters.

F — Extended syntax recognition.

2.3.3 VRM System Compatibility Command

In certain user host computer system configurations the use of ASCII framing characters, DC1, DC2, DC3, or DC4 may perform other system functions. Or, for various reasons, the user may not wish to utilize the ASCII characters ACK, NAK, or even CR or LF. The VRM firmware allows the user the capability to override any one or all of the standard VRM communication protocol characters. (See section 2.3.5.14 for details concerning this command). ASCII control characters vs binary code for the communication protocol characters are shown below.

		U. S. ASCII Code						
ASCII Control Character	ASCII Keyboard	MSB b7	b6	b5	b4	b3	b2	LSB b1
DC1	Q and Control	0	0	1	0	0	0	1
DC2	R and Control	0	0	1	0	0	1	0
DC3	S and Control	0	0	1	0	0	1	1
DC4	T and Control	0	0	1	0	1	0	0
ACK	F and Control	0	0	0	0	1	1	0
NAK	V and Control	0	0	1	0	1	0	1
STX	B and Control	0	0	0	0	0	1	0
CR	M and Control	0	0	0	1	1	0	1
LF	J and Control	0	0	0	1	0	1	0

2.3.4 Command Summary

A summary of the host-to-VRM commands, VRM-to-host and host-to-VRM intermediate messages and VRM-to-host termination/acknowledge messages is given in table 2-1. The table assumes a maximum vocabulary of 100 words; if further assumes that the user has not elected to change

Table 2-1. Command Summary (Sheet 1 of 2)

Action	Host-to-VRM Command	VRM-to-Host Intermediate Messages	Host-to-VRM Intermediate Messages	VRM-to-Host Termination/ Acknowledge Messages
Train (see 2.3.5.1)	DC1,1,xx,yy,bb,T	DC2,8,uu,T		DC2,1,T
Update (see 2.3.5.2)	DC1,2,xx,yy,bb,T	DC2,8,uu,T		DC2,2,T
Reset (see 2.3.5.8)	DC1,3,T			DC2,3,T (if SC6 closed→ R3,T)
Set Reject Threshold (see 2.3.5.5)	DC1,4,zzz,T			DC2,4,T
Port 2 Transfer (see 2.3.5.9)	DC1,5,DATA,T (if F5=0)			DC2,5,T (if F5=0)
Download (see 2.3.5.4)	DC1,6,xx,yy,T	DC2,6,T (if F2=1)	DC3,REFERENCE PATTERN DATA, T	ACK or NAK, T
Upload (see 2.3.5.3)	DC1,7,xx,yy,T	DC4,REFERENCE PATTERN DATA, T		
Select Common Vocabulary (see 2.3.5.7)	DC1,8,xx,yy,T	DC2,9,ww,dd,sss,rr,T (response for speech input) (if F1=1)		DC2,Y,T (if F2=1)
Recognize (selected syntax) (see 2.3.5.6)	DC1,9,xx,yy,T or DC1,F,SYNTAX,T	DC2,9,ww,dd,sss,rr,T (response for speech input) (if F1=1)		DC2,Z,T (if F2=1)
Set/Clear Mode Flags (see 2.3.5.10)	DC1,A,F1,F2, F3,F4,F5,F6,T			DC2,A,T
Read Reject Threshold (see 2.3.5.11)	DC1,B,T			DC2,B,zzz,T
Write Word Boundary Parameters (see 2.3.5.12)	DC1,C,T1,T2, ETHL,MINSM,T			DC2,C,T
Read Word Boundary Parameters (see 2.3.5.13)	DC1,D,T			DC2,D,T1,T2, ETHL,MINSM,T
Change Control Characters (see 2.3.5.14)				

Table 2-1. Command Summary (Sheet 2 of 2)

Symbol	Signification
xx	First word index number.
yy	Last word index number.
bb	Number of training passes.
uu	Training prompt word number
zzz	Reject threshold (three digit).
ww	Recognized winner word number. If ww = ASCII characters FF, a reject has occurred; if ww = ASCII characters LL, then the utterance has exceeded 250 significant samples (F6=1 only).
dd	Difference between winner score and runner-up score. A number greater than 6 indicates the winner is highly probable (99% chance) of being correct.
sss	Winner score. This score must exceed zzz if the winner is not to be rejected.
rr	Runner-up word number. If different scores are small, the correct classification may be the runner-up. The runner-up word number allows the user to observe which words are likely to confuse with one another. In general, the correct classification is nearly always among the top two winners in the vocabulary, i.e., is, the winner and the runner-up.
F1	Sets extended recognize mode flag.
F2	Sets acknowledge-all-host commands including COMMON VOCABULARY, RECOGNIZE, and DOWNLOAD.
F3	Sets high-speed parallel transfer flag.
F4	Sets large-vocabulary mode (future expansion only).
F5	Transmit all noncommand characters from host to Port 2, excluding DC1, DC3, and STX.
F6	Sets word-too-long flag.
T_1	Beginning of word. Absolute spectrum difference of the VRM 16-channel filter bank. Default value = 32.
T_2	Significant spectral difference during a word of the VRM 16-channel filter bank. Default value = 16.
ETHL	Maximum number of insignificant 5 msec samples (silence duration) allowed during an utterance. The default value of ETHL = 32 is an equivalent time interval of 5 msec x32 = 160 msec.
MINSM	Minimum number of significant samples per word. Default value = 16. Assuming all samples are significant this is equivalent to 5 msec x16 = 80 msec.
Y	Command identifier for VRM acknowledge-to-COMMON VOCABULARY command.
Z	Command identifier for VRM acknowledge-to-RECOGNIZE command.
SYNTAX	Word index numbers and delimiters not to exceed 76 characters.

the VRM default framing characters DC1, DC2, DC3 and DC4. The table also assumes the extended recognition flag (F1) and the acknowledge-all-commands mode flag (F2) have been set, and the power-up/reset acknowledge switch, SC-6, has been closed.

2.3.5 Programming Details, VRM/Host Commands

2.3.5.1 TRAIN

Command Sequence:

DC1 1 xx yy bb T

where:

xx* = first vocabulary item number $00 \leq xx \leq 99$
yy* = last vocabulary item number $xx \leq yy \leq 99$
bb = number of training passes $01 \leq bb \leq 64$
T = terminator (CR or CRLF)

The VRM initializes the reference pattern area for classes xx through yy, provides prompt indices for the next expected spoken input, and generates a reference pattern for each vocabulary item. The next word-training prompt index is transmitted to the host by the sequence:

DC2 8 uu T

where:

uu = vocabulary item number $00 \leq uu \leq 99$

At the end of the training process, the VRM transmits DC2 1 T and waits for the next host-to-VRM command.

If the xx = yy, the VRM performs single word training. The number of training passes is set equal to "bb" for the specified word. If "bb" = 0, the VRM utilizes the previous number of training passes for the specified vocabulary item. If no previous training exists, the VRM automatically performs five training passes for the single word, given that bb = 0 with xx = yy.

*The VRM/host commands xx and yy must conform to the limitations of the VRM model selected: 39 for Model 40, and 99 for Model 100.

2.3.5.2 UPDATE

Command Sequence:

DC1 2 xx yy bb T

where:

xx = first vocabulary item number $00 \leq xx \leq 99$
yy = last vocabulary item number $xx \leq yy \leq 99$
bb = number of training passes $01 \leq bb \leq 64$

The VRM UPDATE operation is the same as for the TRAIN mode except that the reference pattern area is not initialized, but retains the previous training set. Additional examples of some or all of the vocabulary items may then be used to update the reference patterns. The next word-training prompt index is transmitted to the host by the sequence:

DC2 8 uu T

where:

uu = vocabulary item number $00 \leq uu \leq 99$

At the end of this operation, DC2 2 T is sent and the VRM awaits the next command. If xx = yy and bb = 00, then the VRM defaults to 1 update pass.

2.3.5.3 UPLOAD

Command Sequence:

DC1 7 xx yy T

where:

xx = first vocabulary item number $00 \leq xx \leq 99$
yy = last vocabulary item number $xx \leq yy \leq 99$

The UPLOAD command requests the VRM to transmit the reference patterns for class xx through class yy to the host for storage or manipulation. The VRM responds with:

DC4 <N Characters of ASCII Data in 4-bit Binary Format >T

where:

N = (1 + yy − xx) 68

If the number of characters, N, received by the host does not equal that requested by the UPLOAD command, the host must reinitiate the UPLOAD command.

If the high-speed parallel mode flag is set (F3=1), then the VRM responds with:

DC4 < N Characters of Data in 8-bit Binary format >T

where:

$$N = (1 + yy - xx)\ 34$$

Note that setting F3=1 halves the number of characters transmitted via the parallel port. For serial data uploading, the Flag F3 should be left to its default value; i.e., F3=0.

With the parallel upload, each character, assuming no host delays, may be uploaded to the host in 25 microseconds. Serial character transfer times depend upon the selected baud rates and are, in general, much slower.

The approximate transfer time for uploading reference patterns is given by the following formulas; assuming no host delays:

$$\text{Serial Transfer Time (seconds)} = \frac{(1 + yy - xx)680}{\text{Baud Rate}}$$

with F3=0

Parallel Transfer Time (milliseconds) = (1 + yy − xx) (1.7)

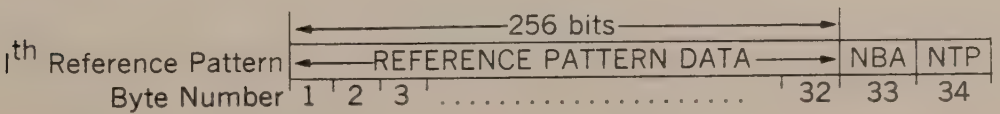
with F3=1

Parallel Transfer Time (milliseconds) = (1 + yy − xx) (0.85)

Several useful manipulations that the host program may perform using uploaded reference patterns are described as techniques to determine which reference patterns, if any, may be causing confusion.

1) The last four ASCII characters of each reference pattern with F3=0 consist of two pairs of numbers called the NBA (number of bits in agreement) and the NTP (number of training passes). The NTP value can be used to locate reference patterns which have been trained more or fewer times relative to other patterns. The recommended number of passes is suggested at seven to ten. If F3=1, the last two bytes contain the binary representation of NBA and NTP respectively.

2) Each reference pattern can be examined as a "master word", I, whose 256 bits (32 bytes) of voice pattern data can be compared to any and/or all other upload patterns, J. The comparison may be acomplished in the following manner:



J^{th} Pattern Being Compared	← REFERENCE PATTERN DATA →						NBA	NTP
	1	2	3	.....		32	33	34

Next I and J are compared by form NBM, that is the total number of ones in agreement in all 256 bits. Analytically this may be expressed by the following formula:

$$NBM\ (I,J) = \sum_{K=1}^{256} \left[BIT_K(I)\right]\ AND\ \left[BIT_K(J)\right]$$

Next a delta score is computed using the following formula:

$$\text{Delta Score } (I,J) = 128 * \left[1 - \left[\frac{(NBM\ (I,J) - 8)}{NBA\ (J)}\right]\right]$$

If the delta score is greater than 6, a user can be very confident that words I and J will not be confused. If the delta score is greater than or equal to 3, fair separability can be expected. A delta score of less than 3 suggests substituting an alternate I or J word or retraining. These delta scores assume 7 to 10 training passes to be statistically significant.

2.3.5.4 DOWNLOAD
Command Sequence:

DC1 6 xx yy T

where:

xx = first vocabulary item number $00 \leq xx \leq 99$
yy = last vocabulary item number $xx \leq yy \leq 99$

If the acknowledge-all mode flag is set (F2=1), then VRM acknowledges the host command via:

DC2 6 T

If F2=0, then the VRM simply awaits the downloading of host data and does not acknowledge the DOWNLOAD command, but responds to the host following receipt of the reference pattern data.

The DOWNLOAD command prepares the VRM to accept the reference patterns for class xx through class yy; for example, to change vocabularies or speakers. The host transmits:

DC3 < N Characters of ASCII Data in 4-bit Binary Format > T

where:

N = (1 + yy − xx) 68

If the high-speed parallel mode flag is set (F3=1), then the VRM expects the host to transmit:

DC3 < N Characters of Data in 8-bit Binary Format >T

where:

N = (1 + yy − xx) 34

The VRM computes N and counts incoming characters. At the end of the transmission, the VRM responds with ACK T, if the count is correct, and NAK T, if the count is incorrect. If the response is NAK T, the host must retransmit the sequence starting with DC3. If the response is ACK T, the VRM awaits the next command. This operation replaces part or all of the previous contents of the VRM memory with the host-transmitted reference patterns.

The approximate transfer time for downloading reference patterns is given by the following formulas. As previously described, the parallel download requires 160 microseconds/character.

$$\text{Serial Transfer Time (seconds)} = \frac{(1 + yy - xx)680}{\text{Baud Rate}}$$

with F3=0

Parallel Transfer Time (milliseconds) = (1 + yy − xx) (3.06)

with F3=1

Parallel Transfer Time (milliseconds) = (1 + yy − xx) (1.53)

At a 9600 serial baud rate, a 100-word transfer requires about 7.1 seconds. Operated in the parallel mode with F3=0 allows downloading of 100 words in 0.306 second, while with F3=1 only 0.153 second is required.

2.3.5.5 SET REJECT THRESHOLD

Command Sequence:

DC1 4 zzz T

where

zzz = no-decision reject threshold $0 \leqslant zzz \leqslant 128$

The no-decision or reject threshold (RTHL) is accepted by the VRM and used during the RECOGNITION mode. After receiving RTHL, the VRM responds with:

DC2 4 T

Recognition in the VRM is accomplished by counting matching bits of the patterns generated by the input utterance with the reference pattern created during training. In matching each reference against the incoming utterance pattern, a normalized score is computed. The maximum value of this score is 128. The reference pattern yielding the highest score relative to the incoming utterance is classified as the vocabulary item just spoken; however, the utterance is rejected if the highest score fails to equal or exceed RTHL. For this reason, a high RTHL setting will cause the VRM to discriminate and reject invalid utterances and/or noise. Although a higher RTHL setting yields better invalid utterance and/or noise discrimination, it can cause rejection of valid utterances if set too high. The "score" on a vocabulary word tends to increase with training, maximizing at 7 to 10 training passes to a value in the range of 118 to 124. RTHL should be set accordingly, with typical values of 100 to 106 for 3 to 5 training passes and up to 110 to 118 for 7 to 10 passes.

An initial conversative estimate frequently used in setting RTHL is:

RTHL = 96 + 2 x (Training Passes)

2.3.5.6 RECOGNITION
Command Sequence:

DC1 9 xx yy T

where:

xx = first vocabulary item number $00 \leq xx \leq 99$
yy = last vocabulary item number $xx \leq yy \leq 99$

If the acknowledge-all-commands mode flag is set (F2=1), then the VRM responds to the host as follows:

DC2 Z T

The host may use the "Z" to serve as an indicator to continue with other processing until the VRM actually receives spoken input.

The extended syntax form of the RECOGNITION command is as follows:

DC1 F xx, gg/hh, ii, jj/yy . . . T

where:

each letter pair are vocabulary item numbers ≤ 99,
gg/hh and jj/yy, are number ranges $gg \leq hh$ and $jj \leq yy$,
and total length of command is ≤ 80 characters.

Using the alternate form permits the host to specify multiple syntax groups, or a range of vocabulary words, which may be classified. Consider, for example, the following application. A control system consists of four valves that can be opened, closed, or checked; three relays that can be opened, closed, or checked; and two gates that can be opened or closed only. "Correction" can be used as a fail-safe word to allow any operation to be changed. "Perform Operation" causes the specified operation to be performed. Part of the vocabulary for this application might be grouped as follows:

Word Number	Vocabulary	Possible Syntax Groupings
.	.	.
.	.	.
.	.	.
41	"One"	Numbers group
42	"Two"	
43	"Three"	
44	"Four"	
45	"Open"	Operation group
46	"Close"	
47	"Check"	
48	"Valve"	Control-type group
49	"Relay"	
50	"Gate"	
51	"Correction"	Error-control group
52	"Perform Operation"	Special action group
.	.	.
.	.	.
.	.	.

The allowable commands to the controls must always be stated as follows:

... {operation group} {control-type group} {numbers group} {special action group} ...

In this way, a requested operation is first verified before being performed. If the verification proves faulty, the request can be changed before controls are changed. Using the alternate form of the recognition command, the host program might specify any of the following combinations, or others, to allow only the correct group of words to be recognized at the proper times:

```
DC1 F 51, 45/47 T      ← allows only "correction" plus operations group
DC1 F 51, 41/44 T      ← allows only "correction" plus numbers group
DC1 95152 T            ← standard form also available
DC1 F 41/44, 51, 15/24 T ← allows some additional words elsewhere in vocabulary
```

In section 2.3.5.7, the COMMON VOCABULARY command is discussed. It is noteworthy that by using the "F" (extended syntax) form, the effects of the COMMON command can be combined

with those of the standard form RECOGNITION command. Thus, any subset of the total vocabulary, regardless of its grouping, can be specified by the host for recognition.

The RECOGNITION command causes the VRM to compare incoming words with the reference patterns for class xx through class yy. The class index with the highest score will be output by the VRM in the format:

DC2 9 ww dd T

where:

ww = winning class index (or FF ASCII, if none exceeded RTHL; or LL ASCII, if an utterance exceeds 250 saved samples and F6=1).

dd = the last two digits of the difference between the winner and the runner-up score, $0 \leq dd \leq 99$.

The quantity "dd" can be used by the host or the user to determine which utterances, if any, are a source of confusion. If a word consistently returns a "dd" quantity that is less than five, the user should consider updating his training or changing the vocabulary word.

If the extended recognition mode flag is set (F1=1), then the VRM response to an input speech utterance is:

DC2 9 ww dd sss rr T

where:

dd = the last two digits of the difference between the winner's and the runner-up's score.
sss = three-digit winner's score.
rr = runner-up word number index (two digits).
ww = winning class index (or "FF" ASCII, if none exceeded RTHL; or "LL" ASCII, if an utterance exceeded 250 saved samples and F6=1).

The value sss may be utilized by the host program to automatically set the reject threshold, RTHL, by monitoring sss when ww = FF. The value rr may be utilized to determine which words are the nearest neighbors when dd is small (less than 5). If the host program is so designed, the words "try again" or other user-selected command utterances may be utilized to choose rr over ww if dd <5 in a closed-loop interactive man/machine feedback manner. This ensures recognition accuracy in excess of 99% even with difficult vocabularies such as the "100 Word" Interstate Electronics Corporation test vocabulary.

The VRM remains in the RECOGNITION mode until the host issues a new control command.

2.3.5.7 COMMON VOCABULARY
Command Sequence:

DC1 8 xx yy T

where:

xx = first vocabulary item number $00 \leq xx \leq 99$
yy = last vocabulary item number $xx \leq yy \leq 99$

The COMMON VOCABULARY command can be issued by the host whenever the VRM is in the RECOGNITION mode. It instructs the VRM to append class xx through class yy to the current vocabulary set as given by the last RECOGNITION command. This remains in effect until a new DC1 8 xx yy T command is received. If xx = yy = 00, the VRM reverts to the current vocabulary as given by the last RECOGNITION command (no common vocabulary). After receiving this command, the VRM remains in the RECOGNITION mode until instructed otherwise.

If the acknowledge-all-command mode flag is set (F2=1), then the VRM acknowledges the host command by:

DC2 Y T

2.3.5.8 RESET
Command sequence:

DC1 3 T

A RESET command causes the VRM to reinitialize all flags, I/O ports, and hardware functions. It must be used if it is desired to change the I/O configuration during VRM operation (i.e., change in data format, serial/parallel operation, handshaking, etc.). This command does not alter the current reference pattern set. Upon completion of the reset operation, the VRM responds with:

DC2 3 T

The same response is transmitted to the host when power is applied to the VRM. In those cases where DC2 cannot be accepted by the host on power-up or reset, putting switch SC-6 to the OPEN position changes DC2 3 T to R 3 T.

This command sets all mode flags to zero, and sets all word boundary parameters (T_1, T_2, ETHL and MINSM) to the default values (32, 16, 32, 16, respectively).

2.3.5.9 Port 2 Transfer

Command Sequence:

DC1 5 (character string) T

where:

Character string = any ASCII characters except the terminator T

After receiving DC1 5, the VRM retransmits all the following characters up to and including the terminator to serial Port 2; they are received as input at host serial Port 1 or at the host parallel data port. This data may be used to control a display, printer, or other peripheral device. Upon completion of the transmission, the VRM sends:

DC2 5 T

The VRM then resumes normal operation.

Any character sent to serial Port 2 by a peripheral causes the VRM to interrupt its current operation and retransmit (or echo) data from Port 2 to Port 1 or the parallel data output port. If the host serial is selected, then Port 1 is the serial echo port. If the host parallel is selected, then the parallel output port is used to echo Port 2 entries.

When the transmit-all mode flag F5=1 is set, or Switch SB-1 is closed, then all host characters which are not part of a host command are relayed to Port 2 and need not be preceded by DC1 5. With F5=1, only special control characters used for framing a VRM command are not transmitted from the host to Port 2, i.e., the characters DC1, DC3 and STX. This feature allows host response to recognized utterances and host prompts to be output to Port 2 with a minimal amount of host coding.

Note: Switch SB-1 "closed" enables Port 2 transfer automatically, overriding software controls, SB-1 "open" allows normal software control of Port 2.

2.3.5.10 Set/Clear Mode Flags

This command allows 6 flags to be set to "1" or cleared to "0". The set or clear is accomplished via ASCII characters "1" or "0". The format for this command is:

DC1 A F1 F2 F3 F4 F5 F6 T

All flags F1 through F6 are cleared by a POWER-UP or RESET command.

The following defines the function of each of the 6 flags:

Flag	Function
F1	Sets the extended recognition data mode.
F2	Sets acknowledge-all-commands including COMMON VOCABULARY.
F3	Sets the high-speed parallel transfer mode.
F4	Sets the large-vocabulary mode. Default value is zero. This flag is reserved for future use and must always be set = 0. A nonzero value for F4 may cause the VRM to malfunction.
F5	Transmits all noncommand characters from host to Port 2, excluding DC1, DC3 and STX.
F6	If an utterance exceeds 250 saved samples, then the word number, ww, in a recognition response, is replaced by the ASCII characters "LL" (utterance too long).

The VRM acknowledges the set/clear mode flag command by:

DC2 A T

and awaits the next host command.

2.3.5.11 Read Reject Threshold

Using verbal commands in the host program, it is possible to increment or decrement the VRM reject threshold. This command allows the host to check the VRM current RTHL setting and thereby determine if it is set to an appropriate value for a particular application.

The command format is:

DC1 B T

The VRM acknowledges the command by:

DC2 B zzz T

where:

zzz = current reject threshold value.

2.3.5.12 Write Word Boundary Parameters
This command format is:

DC1 C T_1 T_2 ETHL MINSM T

where:

T_1 = a two-digit number between the limits of 16 and 64.

T_2 = a two-digit number between the limits of 08 and 64.

ETHL = a two-digit number between the limits of 08 and 64, requiring a silence pause of 40 to 320 msec, respectively.

MINSM = a two digit number between the limits of 16 and 32.

The default values of T_1, T_2, ETHL and MINSM are respectively 32, 16, 32 and 16.

The VRM recognition algorithm determines the beginning of an utterance by detecting an increase in spectral energy relative to that of the silence interval following the previous utterance. In algebraic form we define:

$$D = \sum_{i=1}^{16} \left[f_i(c) - f_i(L) \right]$$

where:

$f_i(c)$ = the current sampled spectral values of each of the 16 VRM filter energy levels.

$f_i(L)$ = last significant sampled spectral value of each of the 16 VRM filter energy levels.

If:

$D \geq T_1$, then begin new word
$D < T_1$, continue search for beginning of word.

An average value of T_1 over the "100 word" test vocabulary is 55; however, T_1 must be selected small enough to detect even the weakest consonants if they are present in the vocabulary. Example of such weak consonants are f, h, m, and n, which at the onset of a new utterance are in the neighborhood of D = 16. However, in the presence of nonstationary background noise it may be desirable to raise T_1 beyond the default value of 32.

The value $D > T_2$ must be exceeded once the beginning of a word, $D > T_1$, occurs, in order for a sample to be saved. If T_2 is made too large (greater than 64), the possibility of collecting insufficient samples to exceed the minimum allowable samples per word (MINSM) of 16 exists. If T_2 is made too small, the number of insignificant samples necessary to delineate the end of a word (ETHL count) may not be satisfied unless a totally quiet environment exists and breath noise is minimal.

It can be seen that the parameters T_1 and T_2 must be set judiciously and validated in the actual operating environment for the vocabulary in question, if system performance is to be optimized.

The parameter ETHL is an end-of-word counter which increments each time $D < T_2$ and resets to zero each time D exceeds T_1. This parameter can be used to accommodate phrases having long pauses (increase ETHL), or to speed up system response for short, quick speech (decrease ETHL).

The VRM acknowledges the write word boundary parameters command with:

DC2, C, T

and awaits the next command.

2.3.5.13 Read Word Boundary Parameter

This command format is:

DC2, D, T

The VRM response is:

DC2, D T_1 T_2 ETHL MINSM T

Where T_1, T_2, ETHL, and MINSM are two-digit ASCII characters having meaning as previously described (2.3.5.12).

2.3.5.14 CHANGE CONTROL CHARACTERS – This single command can alter the framing characters, (i.e., DC1, DC2, DC3, DC4), acknowledge character, nonacknowledge character, or termination character to satisfy most software compatibility problems.

The format of this command is:

STX, FRAMING CHARACTER1, FRAMING CHARACTER2, FRAMING CHARACTER3, FRAMING CHARACTER4 (i.e. DC1, DC2, DC3, DC4), ACK CHARACTER, NAK CHARACTER, TERMINATOR, T

The framing characters, new ACK, new NAK, and new termination characters are exactly one ASCII character.

The VRM response to this command is:

NEW FRAMING CHARACTER2, X, NEW TERMINATOR

This command is typically given upon occurrence of a reset or power-up, following DC2, 3, T (if SC-6 is closed), or R 3 T (if SC-6 is open) as transmitted by the VRM to the host. All seven fields must be specified in the original host-to-VRM command. These fields may be the same as those defaulted to by a reset or may be new characters. The user-selected new framing characters replace any or all VRM standard framing characters; i.e., DC1, DC2, DC3, and DC4.

Upon completion of download, the VRM will respond with the new user-defined ACK or NAK character. Also, the new NAK character is then used by the error code message. The format remains the same as the former message with the NAK character (NEW NAK 3 T). If the new termination character is CR, then the new T=CR; if the new termination character is LF, then the new T=CR, LF. Any other termination character may be selected (for example, ETX) to replace the onboard switch selected character, except a termination character may not be a DC1, DC2, DC3, or DC4 ASCII character.

As the result of the hardware reset, the VRM returns to the default framing, ACK, NAK and the onboard switch set termination characters. As a result of the software reset, the VRM responds with the current framing character, 3, current terminator, and then returns to the default framing characters (DC1, DC2, DC3, DC4), ACK, NAK, and onboard switch setting of terminator character.

If F5=1, then Port 1-to-Port 2 transfers all characters except the new framing character. CHANGE CONTROL CHARACTER command cannot be reissued by the host until a reset (manual or software) or power-up occurs.

2.4 ELECTRICAL AND INTERFACE REQUIREMENTS

2.4.1 Power Requirements

VRM power requirements are summarized as follows:

Voltage ±5% (Vdc)	Current (mA)
+5	1000
+12	120
−12	120

A low-cost power supply suitable for the VRM current and voltage demands is the Power-One Inc., Model HTAA-16W. The dimensions of this supply are 6.5 by 4.0 by 1.62 inches. The HTAA-16 provides at least twice the rated VRM +5 current and approximately three times the ±12 volt current required. With 4K of RAM (100-word vocabulary), the ±5V current demand is 1000 mA; while with 2K of RAM (40 words), the current reduces to 720 mA.

2.4.2 Audio Input
See section 2.1.5 and figure 2-2 for details.

2.4.3 Digital Input/Output

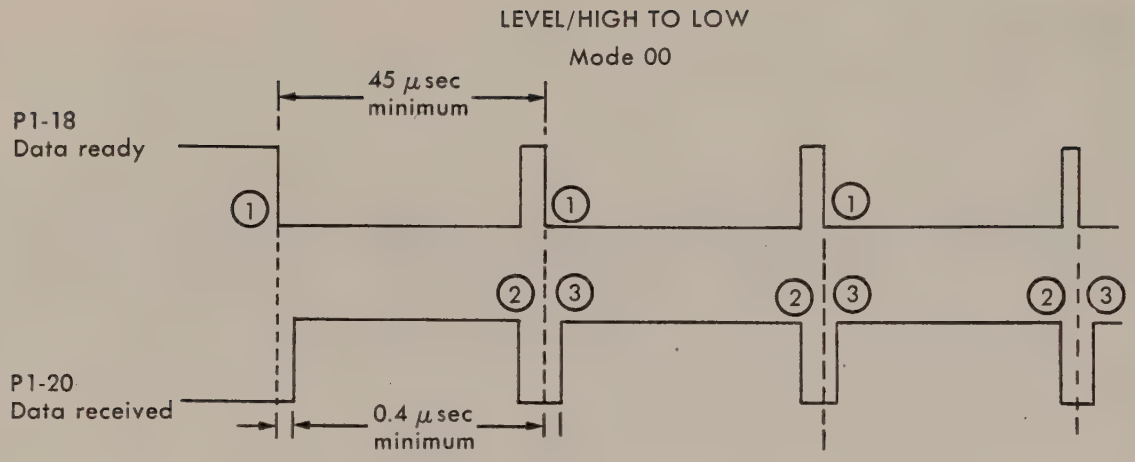
2.4.3.1 Parallel Standard – The standard VRM input or output is via the parallel TTL interface. This interface provides eight data input lines, eight data output lines, and four handshaking control lines. Each output is capable of sinking 1.6 mA current while providing a low voltage (logic 0) output of 0.4 volt or less. The high output (logic 1) voltage is greater than 2.4 volts and capable of sourcing 100 microamps. The TTL input and output should use driver/receiver buffering if the distance from the host to the VRM exceeds 15 feet. A logic "1" input level should be 2.4 volts and capable of delivering 100 microamps. A logic "0" input corresponds to 0.4 volt or less and should be capable of sinking 1.6 mA, (see figure 2-2 for additional details.)

2.4.3.2 Optional Serial I/O – Serial options include one or two serial I/O ports with selectable baud rate, word format, and onboard EIA (RS-232C) or current loop switch-selection option. The RS-232C inputs and outputs provide +12 volts for a logic "0" (space) and −12 volts for a logic "1" (mark). These levels are midrange of the RS232C; i.e., +3 to +25 is the specified logic "0" range, while −3 to −25 is the specified logic "1" range. The request-to-send line, RTS, and clear-to-send line, CTS, are for use with RS-232C levels only. RS-232C communication is useful at distances up to 50 feet.

The current loop interface may be operated at distances up to 2000 feet. The VRM provides a transmitting current source of 20 mA, and the receiving current loop of the VRM shares a common ground return with the host.

2.4.4 Data/Command Synchronization
For parallel port transmissions, the handshaking mode may be selected, as set by switch C, positions 3 and 4, or the equivalent external TTL level applied to P1-25 and P1-27, to allow the VRM four handshaking conventions. These four conventions provide pulse or level with positive or negative transitions as handshaking modes of operation. A parallel input handshaking diagram is shown in figure 2-3, output handshaking in figure 2-4.

For serial port transmissions, the clear-to-send line, CTS, can be used to inhibit VRM transmission, thereby eliminating interrupts or polling until the host is again available. If CTS is caused to be −12 volts (mark), no further VRM transmission occurs; however, the VRM continues in the recognize mode as commanded and stores the word numbers associated with the last few utterances. Because the VRM is always ready to receive data, the RTS line is always +12 volt (space).

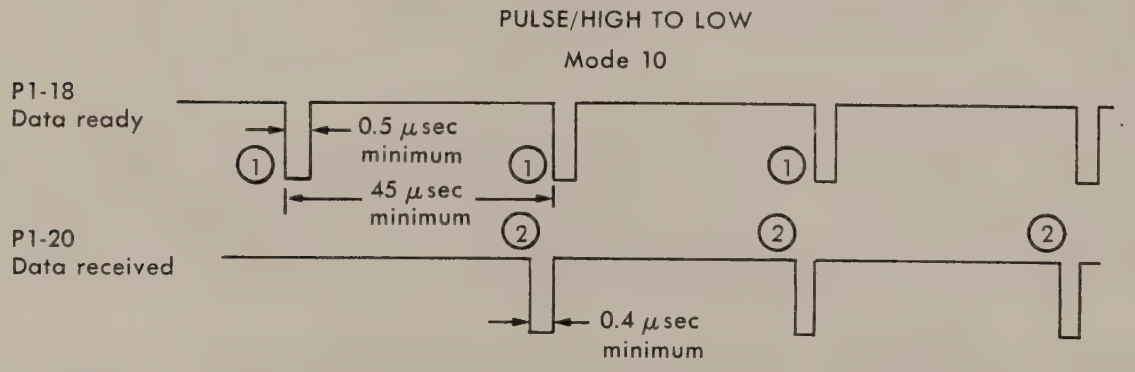


① Host brings data-ready line low to indicate data transmitted; waits for data-received line to go low.

② When data accepted, VRM brings data-received line low; automatically restored to high by next data-ready high-to-low transition, ③.

Note: Assumes no host delays following VRM data received.

Mode 01 identical, except active transition on data-ready line is low-to-high.



① Host sends pulse to indicate data ready and waits for return.

② VRM returns pulse when data accepted/received.

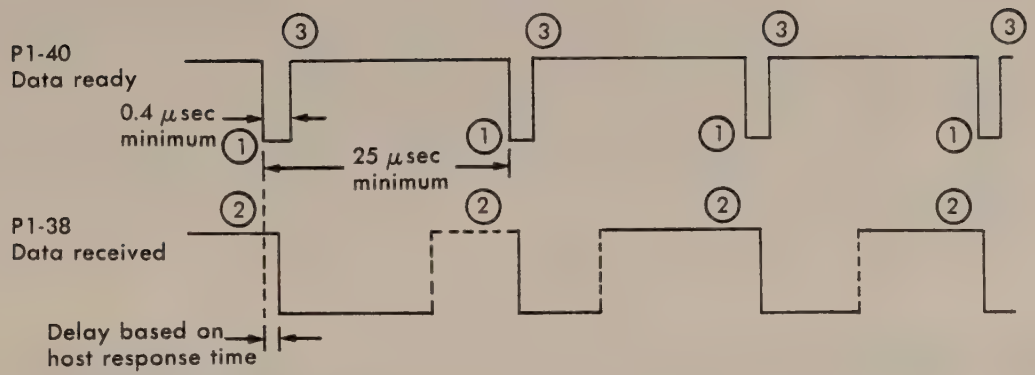
Note: Assumes host reinitiates data ready immediately following VRM data received pulse.

Mode 11 identical, except active transition on data-ready line is low-to-high.

Figure 2-3. Parallel Input Handshaking (Host to VRM)

LEVEL/HIGH TO LOW

Mode 00



(1) VRM brings line low, indicating data ready.

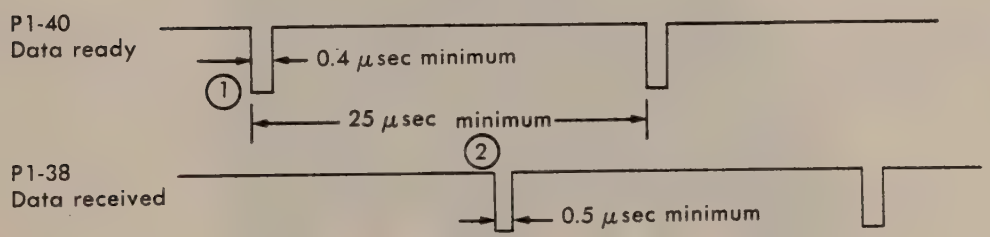
(2) Host brings data-received line low when data accepted, automatically restoring data-ready line high. (3)

Note: Assumes that host accepts data nearly instantaneously, and that data received goes high before next data is ready from VRM.

Mode 01 identical, except active transition of data received line is low-to-high.

PULSE/HIGH TO LOW

Mode 10



(1) VRM sends pulses to indicate data ready; waits for return.

(2) Host returns pulse when data accepted.

Note: Assumes that host data received pulse occurs before next VRM data is ready.

Mode 11 identical, except active transition on data-received line is low-to-high.

Figure 2-4. Parallel Output Handshaking (VRM to Host)

2.4.5 Error Conditions

The VRM tests two error conditions: (1) hard errors and (2) illegal command strings. A hard error occurs if a serial port indicates a receiver overrun, parity, or a framing error. The VRM detects an illegal command string if an incorrect number of characters (eight or more) are transmitted for a given command, if the command identifier field does not contain 1 through 9 or A through F, or if a nondigit character in the bb, xx, yy, zzz field is detected. The VRM does not provide for a timeout. It will wait for a terminator prior to responding. For host messages sent sequentially prior to a VRM response, the VRM makes provision for striking several messages, responding and acknowledging as directed.

Upon detection of any error, the VRM transmits a NAK 3 T to the host.

2.4.6 System Configuration

The VRM accommodates several user system configurations that can be selected by onboard switches and/or external TTL logic levels.* Five switches (see figure 2-5) are located on the board. The VRM system configuration must be set before any host/VRM communication can properly transpire. Switches B and C are read only on a power-up, or a manual or command reset.

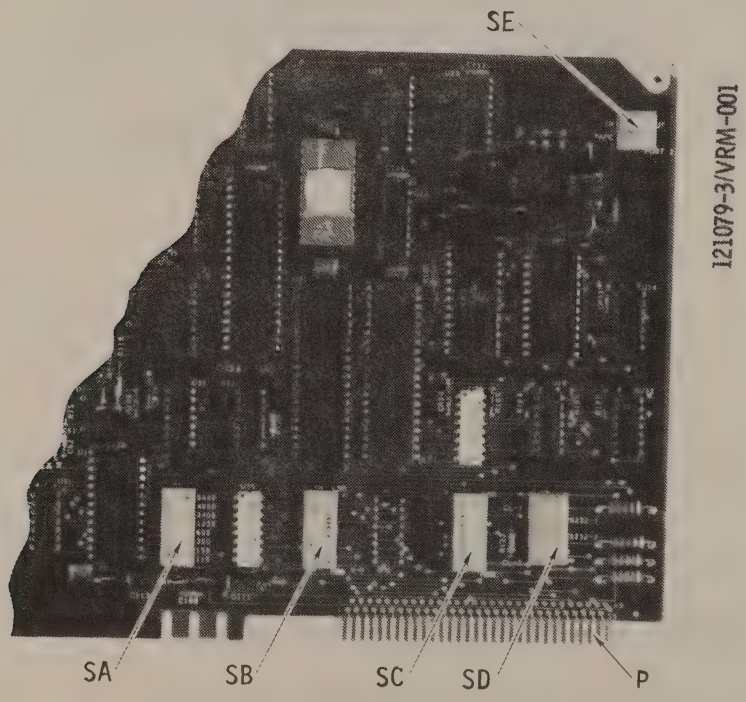


Figure 2-5. VRM Switch Locations

*If TTL logic levels are used instead of onboard functional control with switches SB and SC, these VRM switches must be left in the OPEN position.

The VRM switch labeled SA is used to set the serial baud rate. Switch SB sets the serial word format. The multipurpose switch, SC, is for selection of serial or parallel host communications, termination (i.e., carriage return or carriage return and line feed), echo options, and for selecting the parallel handshaking communication mode and power on/reset acknowledge, as required by the host. Switch SC1 must be left in the CLOSED position to assure proper sampling of the speech input. Use the OPEN position for 4-MHz boards only. The echo option, SC8 open, causes the VRM to ignore the transmitted VRM-to-host message in event the host echoes the VRM output, except for data which originates from input to Port 2 with F5=1 (full duplex host communications). Switch SD sets the board for either the 20-mA current loop or RS-232C serial signal communications, and switch SE is used for utilizing or bypassing the onboard preamplifier.

Table 2-2 summarizes the functions vs switch or TTL logic setting that permit matching the host and VRM. The table lists all applicable VRM switches and/or host control connections.

2.5 MECHANICAL

2.5.1 Board Size

The VRM dimensions are 6.75, by 12.00, by 0.062 inches (standard Intel "Multibus TM" board size). (See figure 2-6).

2.5.2 Connectors

The VRM power connector is designated as P2. Power-pin assignments are summarized in table 2-3. The signal connector is designated P1. Signal/pin assignments for this connector are summarized in table 2-4.

2.5.3 Card Cage

A card cage, Intel ISCB 604, suitable for mounting the VRM with three additional boards, interconnects the 86-pin "Multibus TM" connectors. The total system is readily housed in a 3.5-inch RETMA rack enclosure. Standard Applied Engineering, Inc., Part Number 102498 provides an alternate OEM three-slot card cage without 86-pin interconnection.

Table 2-2. Function vs. Switch/Logic Setting (Sheet 1 of 2)

I/O Selection	P1-29 SC-5	Power Up/Reset Acknowledge	P1-31 SC-6	Terminator	P1-33 SC-7	Host Duplex Mode	P1-35 SC-8
Serial	CLOSED*	DC2	CLOSED	CR	CLOSED	No Echo	CLOSED
Parallel	OPEN	R	OPEN	CR/LF	OPEN	Echo	OPEN

Parallel Interface Handshaking Option	SC-3	SC-4
Level – high to low transition	CLOSED	CLOSED
Level – low to high transition	OPEN	CLOSED
Pulse – high to low transition	CLOSED	OPEN
Pulse – low to high transition	OPEN	OPEN
Data input ready	To VRM	P1-18
Data input received	To host	P1-20
Data output received	To VRM	P1-38
Data output ready	To host	P1-40
Switch SC-1 must be in CLOSED position		

Serial Baud Rate	Switch SA Switch in CLOSED position**
9600	1
4800	2
2400	3
1200	4
600	5
300	6
150	7
110	8

*Closed = On
Open = Off

**Only one switch at a time may be closed on switch SA, and all others must be left in the OPEN position.

Table 2-2. Function vs. Switch/Logic Setting (Sheet 2 of 2)

Serial Data Format Selection

Word Length	Parity	Stop Bits	P1-5 SB-3	P1-7 SB-4	P1-9 SB-5
7	Even	2	C	C	C
7	Odd	2	O	C	C
7	Even	1	C	O	C
7	Odd	1	O	O	C
8	None	2	C	C	O
8	None	1	O	C	O
8	Even	1	C	O	O
8	Odd	1	O	O	O

C = Closed* O = Open

RS-232C/Current Loop Control

Port 1 and Port 2

Switch SD		
Serial Port No.	Input/Output Selection	Switch Label
1	RS-232	C1
1	Current	C2
2	RS-232	C3
2	Current	C4

SB-1	
Auto Port 2 Transfer	C
Software Controlled Port 2 Transfer	O

Switch SE

Function	Switch Label
Microphone preamplifier in	In
Bypass microphone preamplifier	Out
RS-232C logic levels for Port 2	Port 2
TTL logic levels for Port 2	Tape

*C = Closed = On
O = Open = Off

Table 2-3. VRM Power-Pin Assignments

Connector/Pin (no.)	Term	Connector/Pin (no.)	Term
P2-1	GND	P2-2	GND
P2-7	+12 volts	P2-8	+12 volts
P2-75	GND	P2-76	GND
P2-79	−12 volts	P2-80	−12 volts
P2-83	+5 volts	P2-84	+5 volts

Table 2-4. Signal Connections for Host Mode

P1-1	P1-2 Parallel input port (LSB) Bit 0
P1-3	P1-4 Parallel input port Bit 1
P1-5 Data format LSB	P1-6 Parallel input port Bit 2
P1-7 Data format	P1-8 Parallel input port Bit 3
P1-9 Data format MSB	P1-10 Parallel input port Bit 4
P1-11	P1-12 Parallel input port Bit 5
P1-13	P1-14 Parallel input port Bit 6
P1-15	P1-16 Parallel input port (MSB) Bit 7
P1-17	P1-18 Data input ready
P1-19	P1-20 Data input received
P1-21	P1-22 Parallel output port (LSB) Bit 0
P1-23	P1-24 Parallel output port Bit 1
P1-25 Parallel I/O handshaking LSB	P1-26 Parallel output port Bit 2
P1-27 Parallel I/O handshaking MSB	P1-28 Parallel output port Bit 3
P1-29 I/O mode LSB	P1-30 Parallel output port Bit 4
P1-31 I/O mode MSB	P1-32 Parallel output port Bit 5
P1-33 Terminator selection	P1-34 Parallel output port Bit 6
P1-35 Echo	P1-36 Parallel output port (MSB) Bit 7
P1-37 Serial in 1	P1-38 Data output received
P1-39 Serial out 1	P1-40 Data output ready
P1-41 CTS 1	P1-42
P1-43 RTS 1	P1-44 +5 volts
P1-45 Current source No. 1 (20mA)	P1-46 GND
P1-47 Baud rate x 16	P1-48 Word boundary detector
P1-49 Serial in 2	P1-50 Reset
P1-51 Serial out 2	P1-52
P1-53 CTS 2	P1-54 +12 volts
P1-55 RTS 2	P1-56 −12 volts
P1-57 Current source No. 2 (20 mA)	P1-58 MIC in sig
P1-59 GND	P1-60 MIC in GND

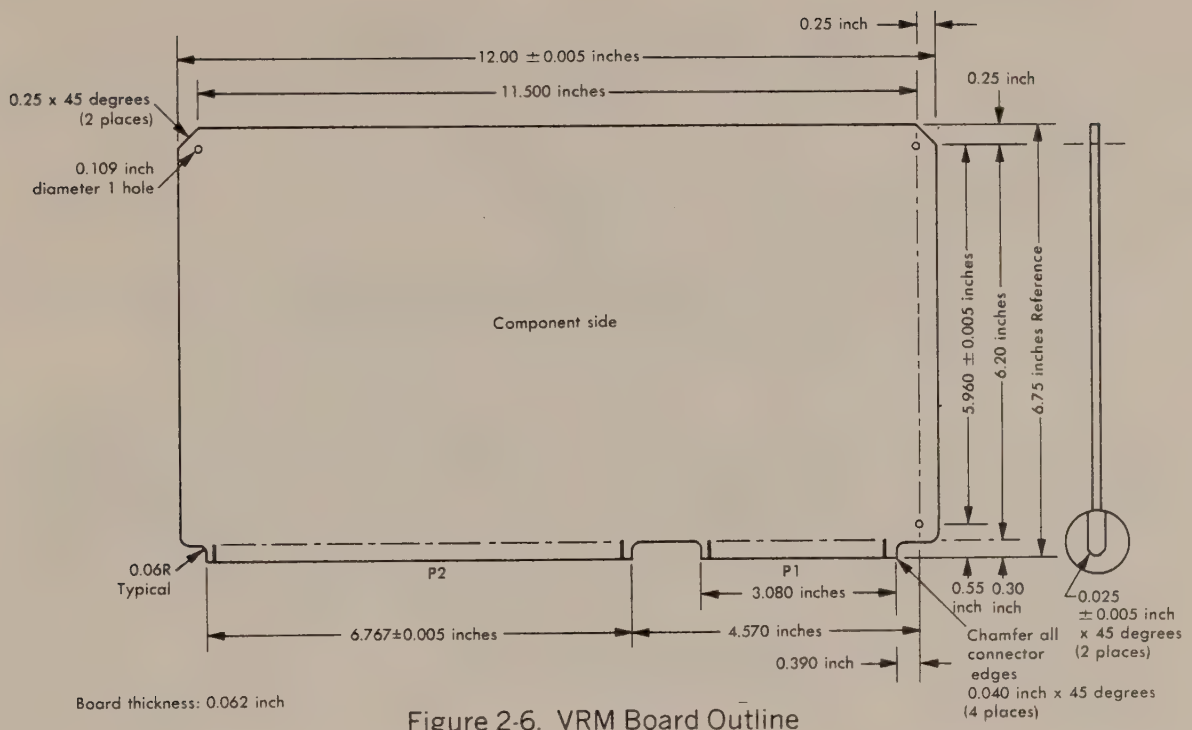


Figure 2-6. VRM Board Outline

Table 2-5. VRM Specification Summary

PERFORMANCE

Vocabulary Size: 40, or 100 isolated words and/or phrases.
Recognition Accuracy: 99+ percent.
Reject Threshold: User selectable.
Longest Utterance Duration: 1.25 second.
Minimum Between-word Pauses: 160 milliseconds. (User-programmable: 40 to 320 milliseconds)
Minimum Word Length: 80 milliseconds. (User-programmable: 80 to 160 milliseconds)
Processing Time: (25+N) milliseconds, where N = active vocabulary size.

HOST COMMANDS

(1) TRAIN, (2) UPDATE, (3) RESET, (4) SET REJECT THRESHOLD, (5) PORT 2 TRANSFER, (6) DOWNLOAD, (7) UPLOAD, (8) COMMON VOCABULARY, (9) RECOGNITION – (A) SET/CLEAR FLAGS, (B) READ REJECT THRESHOLD, (C) WRITE WORD BOUNDARY PARAMETERS, (D) READ-WORD BOUNDARY PARAMETER, (F) EXTENDED SYNTAX, AND (STX) CHANGE CONTROL CHARACTERS.

*SET/CLEAR FLAGS provides six independent user flag options.

DIGITAL INPUT-OUTPUT

Standard:	Parallel TTL input-output, eight data input bits, eight data output bits with four control lines. Data transfer rate: host-to-VRM, 160 microsecond/character, VRM-to-host, 120 microsecond/character.
Optional:	Asynchronous serial interface to host in place of standard parallel interface (Port 1)(110 to 9600 Baud).
Optional:	Asynchronous serial auxiliary interface (useful with peripheral such as a CRT, printer, or display); (Port 2) (Baud rate dependent on Port 1 setting).

MECHANICAL

Card Size:	6.75 by 12.00 by 0.062 inches (standard Intel "Multibus™" card size).

CONNECTORS

Power:	86-pin, 0.156-inch spacing, Viking 2CH43/1AN or equivalent.
Signals:	60-pin, 0.100 inch spacing, AMP PE5 14559 connector (signal connector furnished with board).

ELECTRICAL

Power Requirements: +5 Vdc at 1000 mA, +12 Vdc at 150 mA, −12 Vdc at 150 mA.

Chapter 3

VOTERM HARDWARE DESCRIPTION

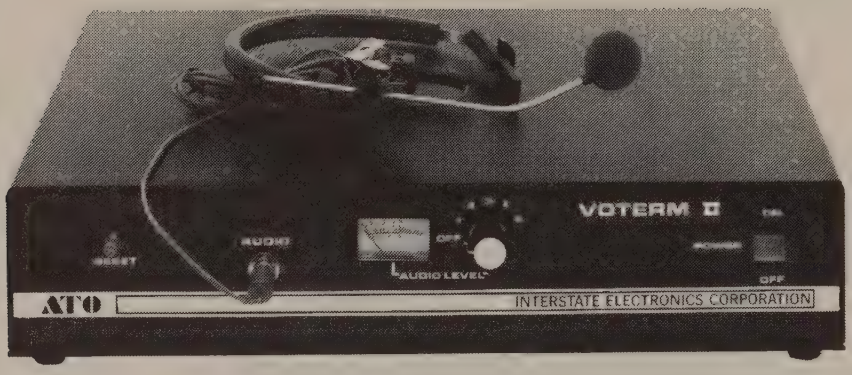
The Voterm is a portable chassis designed specifically to house a VRM and permit easy interconnection of the board with various mini/micro host processors. The chassis comes complete with power supply, fuse (1/2 A), audio indicator LED, Vu meter, audio amplifier, audio level switch, fan, power cord, and a microphone (Shure Brothers SM-10) for use with the VRM.

The front panel (figure 3-1) of the Voterm contains a power on/off switch, audio level switch, microphone audio input jack, VRM manual-reset momentary switch, and one LED indicator. The LED labeled "audio" lights when the VRM detects a spoken word or phrase. The "audio level" Vu meter and switch provide the user a with a means of setting speech-signal amplitude to match normal speaking level. The Voterm Microphone Amplifier PCB (figure 3-2) has a selectable gain, integrated-circuit microphone amplifier. The audio level is monitored by a front-panel meter. Five gain levels provide a wide variety of user speaking intensities compensating for mouth-to-microphone coupling distance and background noise. For normal operation with average speaker, a gain setting of "3" is recommended. On the VRM board, switch "SE" disables the VRM microphone amplifier with the "In-Out" portion of the switch. When using the Voterm, this switch must be in the "Out" position.

The Voterm rear panel (figure 3-3) has four connectors. Two connectors provide serial asynchronous ASCII communications ports: one port for host interfacing, and the other port for an auxiliary peripheral device. The other two connectors provide two parallel I/O ports; one for data, and one for configuration/control. The Voterm inner rear panel contains a slide switch allowing either 230 Vac, 50 cycle; or 115 Vac, 60 cycle operation. Set switch as applicable (see figure 3-3).

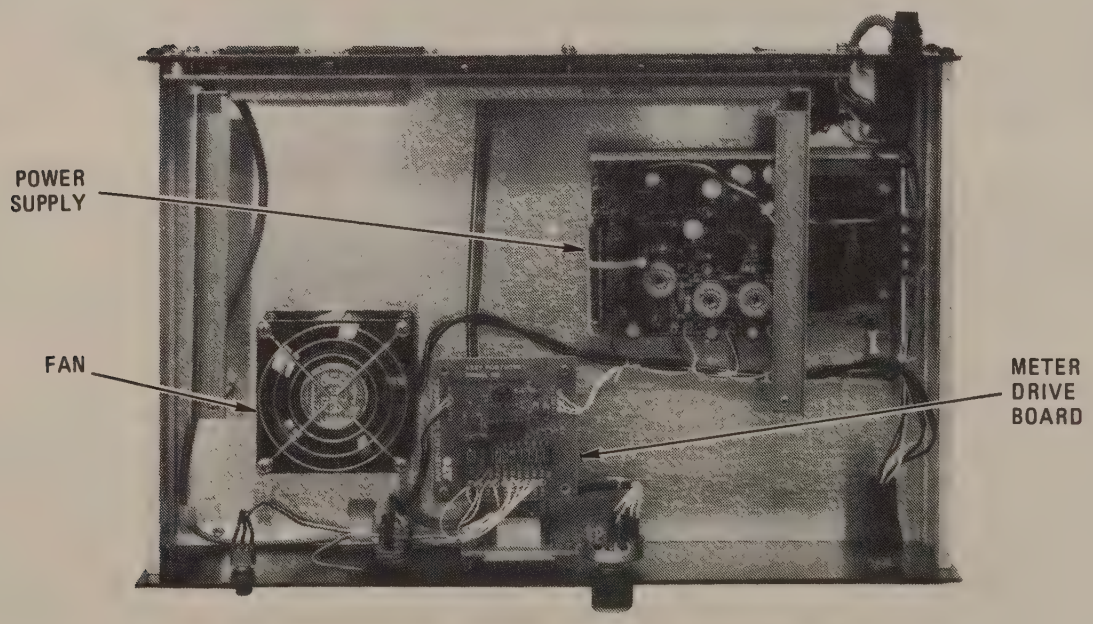
3-1. INSTALLING A VRM

The VRM board is easily installed in the chassis by first removing the four Phillips-head screws on the base of the Voterm case. Next, the chassis is removed from the case by sliding the chassis forward and upward while applying pressure to the rear panel. The VRM is then inserted into the Voterm connector and card guides. The VRM board switches must be set to conform to the user's specific application. Finally, the Voterm chassis is reinserted into the case and the four screws are reinstalled in the base. (Note: care should be taken, before reinserting the chassis, to make certain that the VRM is seated properly.) The Voterm is ready for use as soon as the power cord is plugged into a wall outlet and the front-panel switch is turned on.



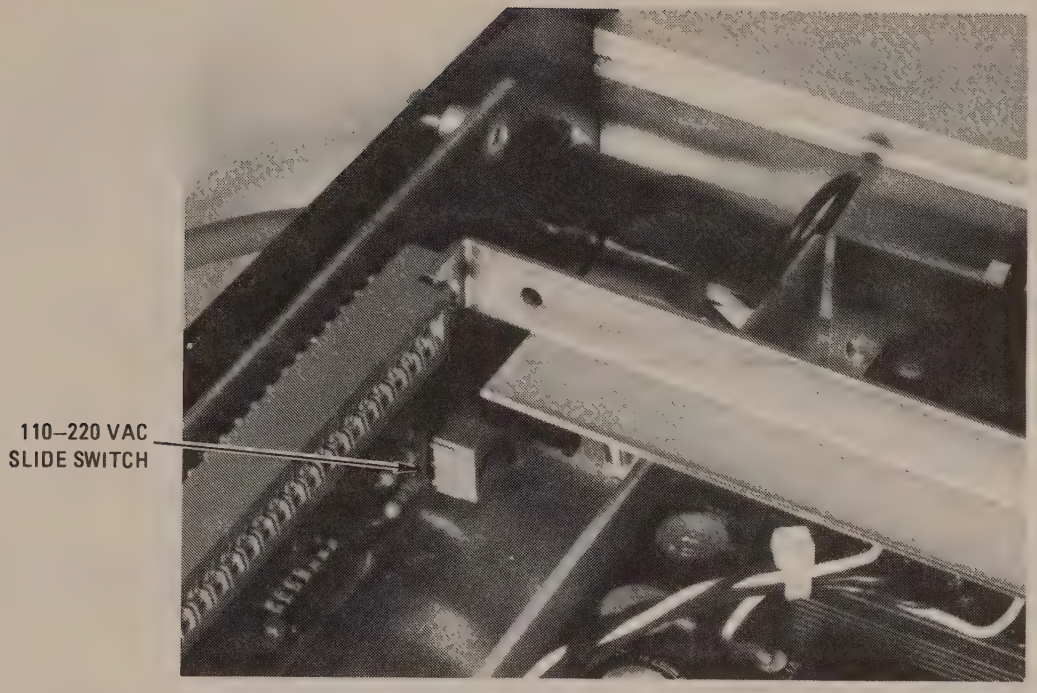
032081-1

Figure 3-1. Front Panel of Voterm II (With Shure SM-10 Microphone)



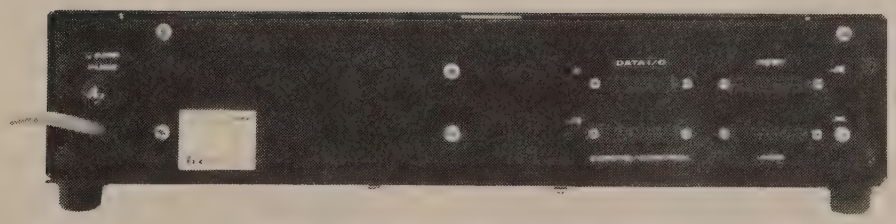
032081-3

Figure 3-2. Top View of Voterm II, Showing Meter Drive PCB



032081–4

Figure 3-3. Top View of Voterm II, Showing 110-220 Vac Power Switch



032081–2

Figure 3-4. Rear Panel of Voterm II

3-2. RESET SWITCH

This front-panel momentary switch enables the user to manually reset the VRM without host-processor intervention. When the VRM is reset, all flags are initialized and all switch/TTL logic configuration and control signals are read. Upon completion of the reset function, the VRM alerts the host processor by transmitting DC2, 3, and a terminator (i.e., carriage return, or carriage return and line feed) or R3 and terminator if SC-6 is in OPEN position.

3.3 CONNECTOR PIN ASSIGNMENTS

The rear panel of the Voterm contains four DB-25S connectors configured as follows:

Tables 3-1 through 3-4 list the pin functions and pin numbers for connectors J1 through J4: tables 3-1 and 3-2 summarize the parallel-port connections, and table 3-3 and 3-4 summarize the serial-port connections. Table 3-5 lists the available power from the Voterm power supply.

3.4 AIR FILTER MAINTENANCE

The slimline filter is of compact, one piece design, capable of filtering dust particles from the atmosphere at all static pressures.

The filter, made of a special aluminum mesh, is alodined and treated with atomized oil. It may be removed easily for cleaning.

CAUTION

Disconnect power cord (unplug unit) before starting to inspect or clean the filter.

3.4.1 Visual Inspection

Check filter every 90 days or sooner depending on conditions and quality of the environment. Ten percent minimal airflow restriction is cause to remove and clean. Do not attempt to clean filter without first removing it from the containing unit.

3.4.2 Removing Filter

Remove the four Phillips-head 8-32 x 5/16 LG mounting screws on bottom of case. Pull the chassis assembly forward carefully until it clears the case. Remove the Phillips-head 8-32 x 3/8 LG mounting hardware from the four corners of the air filter.

3.4.3 Filter Assembly

To assemble the filter, reverse the above operation (paragraph 3.4.2)

3.4.4 Cleaning the Filter

The filter can be easily washed in solvent or vapor degreaser, or cleaned ultrasonically. Air dry. Respray with atomized oil.

Table 3-1. Parallel-Port Connections for Input/Output

Connector/ Pin (no.)	Function
J1-1	Data in (LSB)
J1-2	Data in
J1-3	Data in
J1-4	Data in
J1-5	Data in
J1-6	Data in
J1-7	Data in
J1-8	Data in (MSB)
J1-9	Data in ready
J1-10	Data in received
J1-11	
J1-12	
J1-13	GND
J1-14	Data out (LSB)
J1-15	Data out
J1-16	Data out
J1-17	Data out
J1-18	Data out
J1-19	Data out
J1-20	Data out
J1-21	Data out (MSB)
J1-22	Data out received
J1-23	Strobe (parallel data ready)
J1-24	
J1-25	GND

Table 3-2. Parallel-Port Connections for Configuration/Control

Connector/ Pin (no.)	Function
J3-1	Port-2 transfer mode
J3-2	
J3-3	Serial data format (LSB)
J3-4	Serial data format
J3-5	Serial data format (MSB)
J3-6	
J3-7	
J3-8	
J3-9	
J3-10	Word boundary detector
J3-11	
J3-12	
J3-13	GND
J3-14	
J3-15	
J3-16	Parallel handshaking mode (LSB)
J3-17	Parallel handshaking mode (MSB)
J3-18	Serial parallel
J3-19	Host/standalone mode
J3-20	Terminator character (CR/CRLF)
J3-21	Echo mode (half/full duplex)
J3-22	
J3-23	
J3-24	
J3-25	GND

Table 3-3. Serial-Port Connections to Host Processor

Connector/ Pin (no.)	Function
J2-1	GND
J2-2	Serial in
J2-3	Serial out
J2-4	CTS
J2-5	RTS
J2-7	GND
J2-23	Current loop source

Table 3-4. Serial-Port Connections to Peripheral Device

Connector/ Pin (no.)	Function
J4-1	GND
J4-2	Serial in
J4-3	Serial out
J4-4	CTS
J4-5	RTS
J4-7	GND
J4-23	Current loop source
J4-9	Serial in TTL
J4-11	Serial out TTL

Table 3-5. Available Voterm DC Power

Voltage ±5% (VDC)	Current (mA)
+ 5	1000
+12	280
−12	280

Table 3-3. Serial Port Connector
Host Processor

Connector Pin (J2)	Function
J2-1	GND
J2-2	Serial In
J2-3	Serial Out
J2-4	CTS
J2-5	RTS
J2-7	GND
J2-25	Current [illegible] 0.3

Table 3-4. Serial Port Connector
Peripheral Device

Connector Pin (no)	Function
J4-1	GND
J4-2	Serial In
J4-3	Serial Out
J4-4	CTS
J4-5	RTS
J4-7	GND
J4-25	Current loop s[illegible]ce
[illegible]	Serial In TTL
[illegible]	Serial out TTL

Table 2-5. Available Interim DC Power

Voltage ±5% (VDC)	Current (mA)

INTERSTATE
ELECTRONICS
CORPORATION
SUBSIDIARY OF ATO

1001 E. Ball Road, P.O. Box 3117, Anaheim, California 92803
Telephone 714/635-7210 TWX 910-591-1197 Telex 655443 & 655419
Call toll-free: in the continental U.S. 800/854-6979; in California 800/422-4580

Printed in U.S.A.

Copyright © 1980 Interstate Electronics Corporation 0.5M 4/81